



TITLE:

# Approximation algorithms for the sex-equal stable marriage problem

AUTHOR(S):

Iwama, Kazuo; Miyazaki, Shuichi; Yanagisawa, Hiroki

---

CITATION:

Iwama, Kazuo ...[et al]. Approximation algorithms for the sex-equal stable marriage problem. ACM Transactions on Algorithms 2010, 7(1): 2.

ISSUE DATE:

2010-11-1

URL:

<http://hdl.handle.net/2433/226949>

RIGHT:

© ACM, 2010. This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in 'ACM Transactions on Algorithms' Volume 7 Issue 1, November 2010, <http://doi.acm.org/10.1145/1868237.1868239>; この論文は出版社版ではありません。引用の際には出版社版をご確認ください。 ; This is not the published version. Please cite only the published version.

# Approximation Algorithms for the Sex-Equal Stable Marriage Problem

KAZUO IWAMA and SHUICHI MIYAZAKI

Kyoto University

and

HIROKI YANAGISAWA

IBM Tokyo Research Laboratory

---

The stable marriage problem is a classical matching problem introduced by Gale and Shapley. It is known that for any instance, there exists a solution, and there is a polynomial time algorithm to find one. However, the matching obtained by this algorithm is man-optimal, that is, the matching is favorable for men but unfavorable for women, (or, if we exchange the roles of men and women, the resulting matching is woman-optimal). The sex-equal stable marriage problem, posed by Gusfield and Irving, seeks a stable matching “fair” for both genders. Specifically it seeks a stable matching with the property that the sum of the men’s scores is as close as possible to that of the women’s. This problem is known to be strongly NP-hard.

In this paper, we give a polynomial time algorithm for finding a *near optimal* solution for the sex-equal stable marriage problem. Furthermore, we consider the problem of optimizing an additional criterion: among stable matchings that are near optimal in terms of the sex-equality, find a minimum egalitarian stable matching. We show that this problem is strongly NP-hard, and give a polynomial time algorithm whose approximation ratio is less than two.

Categories and Subject Descriptors: F.2.2 [Analysis of Algorithms and Problem Complexity]: Nonnumerical Algorithms and Problems; G.2.1 [Discrete Mathematics]: Combinatorics—Combinatorial algorithms

General Terms: Algorithms, Theory

Additional Key Words and Phrases: Approximation algorithms, the stable marriage problem, the sex-equal stable marriage problem

---

---

A preliminary version of this article appeared in the 2007 *Proceedings of the 10th Workshop on Algorithms and Data Structures (WADS)*.

This work was supported in part by KAKENHI (16092101, 16092215, 16300002, 17700015, 19200001, 20700009).

Authors’ addresses: K. Iwama, Graduate School of Informatics, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto 606–8501, Japan, e-mail: [iwama@kuis.kyoto-u.ac.jp](mailto:iwama@kuis.kyoto-u.ac.jp); S. Miyazaki, Academic Center for Computing and Media Studies, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto 606–8501, Japan, e-mail: [shuichi@media.kyoto-u.ac.jp](mailto:shuichi@media.kyoto-u.ac.jp); H. Yanagisawa, IBM Tokyo Research Laboratory, 1623-14 Shimotsuruma, Yamato-shi, Kanagawa-ken, 242-8502 Japan, e-mail: [yanagis@jp.ibm.com](mailto:yanagis@jp.ibm.com).

Permission to make digital/hard copy of all or part of this material without fee for personal or classroom use provided that the copies are not made or distributed for profit or commercial advantage, the ACM copyright/server notice, the title of the publication, and its date appear, and notice is given that copying is by permission of the ACM, Inc. To copy otherwise, to republish, to post on servers, or to redistribute to lists requires prior specific permission and/or a fee.

© 2009 ACM 1529-3785/2009/0700-0001 \$5.00

## 1. INTRODUCTION

An instance  $I$  of the stable marriage problem consists of  $n$  men,  $n$  women, and each person's preference list. A preference list is a totally ordered list including all members of the opposite sex depending on his or her preferences. For a matching  $M$  between men and women, a pair of a man  $m$  and a woman  $w$  is called a *blocking pair* if both prefer each other to their current partners. A matching with no blocking pair is called *stable*. Gale and Shapley showed that every instance admits at least one stable matching, and proposed a linear time algorithm to find one, which is known as the Gale-Shapley algorithm [Gale and Shapley 1962]. However, in general, there are many different stable matchings for a single instance, and the Gale-Shapley algorithm finds only one of them (*man-optimal* or *woman-optimal*) with an extreme property: In the man-optimal stable matching, each man is matched with his best possible partner, while each woman gets her worst possible partner, among all stable matchings. Hence, it is natural to try to obtain a matching which is not only stable but also “good” in some criterion.

There are three major optimization criteria for the quality of stable matchings. Let  $p_m(w)$  ( $p_w(m)$ , respectively) denote the position of woman  $w$  in man  $m$ 's preference list (the position of man  $m$  in woman  $w$ 's preference list, respectively). For a stable matching  $M$ , define the *regret cost*  $r(M)$  to be

$$r(M) = \max_{(m,w) \in M} \max\{p_m(w), p_w(m)\}.$$

Also, define the *egalitarian cost*  $c(M)$  to be

$$c(M) = \sum_{(m,w) \in M} p_m(w) + \sum_{(m,w) \in M} p_w(m),$$

and the *sex-equality cost*  $d(M)$  to be

$$d(M) = \sum_{(m,w) \in M} p_m(w) - \sum_{(m,w) \in M} p_w(m).$$

The *minimum regret stable marriage problem* (the *minimum egalitarian stable marriage problem* and the *sex-equal stable marriage problem*, respectively) is to find a stable matching  $M$  with minimum  $r(M)$  ( $c(M)$  and  $|d(M)|$ , respectively) [Gusfield and Irving 1989]. Note that the number of stable matchings for one instance grows exponentially in general (see [Irving and Leather 1986], e.g.). Nevertheless, for the first two problems, Gusfield [Gusfield 1987] and Irving, Leather and Gusfield [Irving et al. 1987] proposed  $O(n^2)$  and  $O(n^4)$  time algorithms, respectively, by exploiting a lattice structure which is of polynomial size but contains the information for all of the stable matchings. Later, Feder [Feder 1992; 1994] improved the complexity of the algorithm for the minimum egalitarian stable marriage problem to  $O(n^3)$ .

In contrast, it is hard to obtain a sex-equal stable matching. The question of its complexity was posed by Gusfield and Irving [Gusfield and Irving 1989], and was later proved to be strongly NP-hard by Kato [Kato 1993]. Thus, the next step should be its approximability for which we have no knowledge so far.

**Our Contributions.** Algorithm  $A$  is said to be a  $c$ -approximation algorithm if  $A(I)/OPT(I) \leq c$  holds for any input  $I$ , where  $A(I)$  and  $OPT(I)$  are the costs

of  $A$ 's solution and an optimal solution, respectively. However, since the optimal cost of the sex-equal stable marriage problem can be zero, this measure cannot be used here. Instead, we consider finding *near optimal* solutions. Let  $M_0$  and  $M_z$  be the man-optimal and the woman-optimal stable matchings, respectively. For a matching  $M$  and a person  $p$ , denote  $M(p)$  the partner of  $p$  in  $M$ . Note that for every man  $m$ ,  $p_m(M(m))$  is minimum when  $M = M_0$  (under the condition that  $M$  is stable), and is maximum when  $M = M_z$ . On the other hand, for every woman  $w$ ,  $p_w(M(w))$  is maximum when  $M = M_0$ , and is minimum when  $M = M_z$ . Thus  $d(M_0) \leq d(M) \leq d(M_z)$  for any stable matching  $M$  (see Fig. 1). Our goal is to obtain a stable matching  $M$  such that  $-\epsilon\Delta \leq d(M) \leq \epsilon\Delta$  for a given constant  $\epsilon$ , where  $\Delta = \min\{|d(M_0)|, |d(M_z)|\}$ . We define the following problem called *Near SexEqual* (NSE). Given a stable marriage instance  $I$  and a positive constant  $\epsilon$ , it seeks a stable matching  $M$  such that  $|d(M)| \leq \epsilon\Delta$  if such  $M$  exists, or answers “None” otherwise. We give a polynomial time algorithm for NSE, which runs in time  $O(n^{3+\frac{1}{\epsilon}})$ .



Fig. 1. The sex-equality costs of stable matchings

NSE seeks an *arbitrary* stable matching whose sex-equality cost lies within some range. However, we may want to find a good one if there are several solutions in the range. In fact, there is an instance  $I$  that has two stable matchings  $M$  and  $M'$  such that  $d(M) = d(M') = 0$  but  $c(M) \ll c(M')$  (see Sec. 4). This motivates us to consider the following corresponding optimization problem *Minimum Egalitarian Sex-Equal stable marriage problem* (MinESE): Given a stable marriage instance  $I$  and a positive constant  $\epsilon$ , find a stable matching  $M$  that minimizes  $c(M)$  under the condition that  $|d(M)| \leq \epsilon\Delta$ , (or the answer “None” if none exists). We show that MinESE is strongly NP-hard, and give a polynomial time  $(2 - (\epsilon - \delta)/(2 + 3\epsilon))$ -approximation algorithm for an arbitrary  $\delta$  such that  $0 < \delta < \epsilon$ , whose running time is  $O(n^{3+2(\frac{1+\epsilon}{\delta})})$ .

Although the details are omitted, our results in this paper can be easily extended to the *weighted* versions of these problems, in which  $p_m(w)$  ( $p_w(m)$ , respectively) represents not simply the rank of  $w$  in  $m$ 's preference list, but an arbitrary score of  $m$  for  $w$  (of  $w$  for  $m$ ), where  $p_m(w) > 0$  ( $p_w(m) > 0$ ) and  $p_m(w) < p_m(w')$  if and only if  $m$  prefers  $w$  to  $w'$  ( $p_w(m) < p_w(m')$  if and only if  $w$  prefers  $m$  to  $m'$ ) for all  $m$  and  $w$ .

**Related Results.** As mentioned above, the minimum regret stable marriage problem and the minimum egalitarian stable marriage problem can be solved in polynomial time [Gusfield 1987; Irving et al. 1987; Gusfield and Irving 1989], but the sex-equal stable marriage problem is strongly NP-hard [Kato 1993]. Romero-Medina [Romero-Medina 2001] provides an (exponential-time) algorithm for finding an optimal solution for the sex-equal stable marriage problem. If we allow ties in preference lists, all these problems become hard, even to approximate, if we seek an optimal *weakly* stable matching: For each problem, there exists a positive constant  $\delta$

such that there is no polynomial-time approximation algorithm with approximation ratio  $\delta n$  unless  $P=NP$  [Halldórsson et al. 2003].

## 2. ROTATION POSET

In this section, we explain a *rotation poset* (*partially-ordered set*), originally defined in [Irving and Leather 1986], which is an underlying structure of stable matchings. Here, we give only a brief sketch necessary for understanding the algorithms given later. Readers can refer to [Gusfield and Irving 1989] for further details.

We fix an instance  $I$ . Let  $M$  be a stable matching for  $I$ . For each such  $M$ , we can associate a *reduced list*, which is obtained from the original preference lists by removing entries. One property of the reduced list associated with  $M$  is that, in  $M$ , each man is matched with the first woman in the reduced list, and each woman is matched with the last man. A *rotation exposed in  $M$*  is an ordered list of pairs  $\rho = (m_0, w_0), (m_1, w_1), \dots, (m_{r-1}, w_{r-1})$  such that, for every  $i$  ( $0 \leq i \leq r-1$ ),  $m_i$  and  $w_i$  are matched in  $M$ , and  $w_{i+1}$  is at the second position in  $m_i$ 's reduced list, where  $i+1$  is taken modulo  $r$ . There exists at least one rotation for any stable matching except for the woman-optimal stable matching  $M_z$ .

For a stable matching  $M$  and a rotation  $\rho = (m_0, w_0), (m_1, w_1), \dots, (m_{r-1}, w_{r-1})$  exposed in  $M$ , *eliminating  $\rho$  from  $M$*  means to move  $m_i$ 's partner from  $w_i$  down to  $w_{i+1}$  for each  $i$  ( $0 \leq i \leq r-1$ ), (and to update a reduced list accordingly). Note that by eliminating a rotation, men become worse off while women become better off. The resulting matching is denoted by  $M/\rho$ . It is well known that  $M/\rho$  is also stable for  $I$ . If a rotation  $\sigma$  is exposed in  $M/\rho$ , then we can similarly obtain another stable matching by eliminating  $\sigma$ .

Now, let  $\mathcal{M}$  be the set of all stable matchings for  $I$ , and  $\Pi$  be the set of rotations  $\rho$  such that  $\rho$  is exposed in some stable matching in  $\mathcal{M}$ . Then, it is known that  $|\Pi| \leq n^2$ . The *rotation poset*  $(\Pi, \prec)$ , which is uniquely determined for instance  $I$ , is the set  $\Pi$  with a partial order  $\prec$  defined for elements in  $\Pi$ . For two rotations  $\rho_1$  and  $\rho_2$  in  $\Pi$ ,  $\rho_1 \prec \rho_2$  intuitively means that  $\rho_1$  must be eliminated before  $\rho_2$ , or  $\rho_2$  is never exposed until  $\rho_1$  is eliminated. It is known that a compact representation of the rotation poset can be constructed in  $O(n^2)$  time.

A *closed subset*  $R$  of the rotation poset  $(\Pi, \prec)$  is a subset of  $\Pi$  such that if  $\rho \in R$  and  $\rho' \prec \rho$  then  $\rho' \in R$ . There is a one-to-one correspondence between  $\mathcal{M}$  and the set of closed subsets of  $(\Pi, \prec)$ : Let  $R$  be a closed subset. Starting from the man-optimal stable matching  $M_0$ , if we eliminate all rotations in  $R$  successively in any order following  $\prec$ , then we can obtain a stable matching. Conversely, any stable matching can be obtained by this procedure for some closed subset. We denote the stable matching corresponding to a closed subset  $R$  by  $M_R$ . For simplicity, we sometimes write  $c(R)$  and  $d(R)$  instead of  $c(M_R)$  and  $d(M_R)$ , respectively. In particular, the empty subset corresponds to the man-optimal stable matching  $M_0$ , and the set  $\Pi$  itself corresponds to the woman-optimal stable matching  $M_z$ . From  $M_0$ , if we eliminate all rotations according to the order  $\prec$ , then we eventually reach  $M_z$ .

For a rotation  $\rho = (m_0, w_0), (m_1, w_1), \dots, (m_{r-1}, w_{r-1})$ , we define  $w_c(\rho)$  and  $w_d(\rho)$ , which represent the cost changes with respect to the egalitarian and sex-

equality measures, respectively, by eliminating  $\rho$ :

$$w_c(\rho) = \sum_{i=0}^{r-1} (p_{m_i}(w_{i+1}) - p_{m_i}(w_i)) + \sum_{i=0}^{r-1} (p_{w_i}(m_{i-1}) - p_{w_i}(m_i)),$$

$$w_d(\rho) = \sum_{i=0}^{r-1} (p_{m_i}(w_{i+1}) - p_{m_i}(w_i)) - \sum_{i=0}^{r-1} (p_{w_i}(m_{i-1}) - p_{w_i}(m_i)).$$

Here, note that  $w_d(\rho) > 0$  for all  $\rho$  since by eliminating a rotation, some men become worse off, some women become better off, and other people remain matched with the same partners. Now, let  $\rho$  be a rotation exposed in a stable matching  $M$ . Then, it is obvious from the definition that  $c(M/\rho) = c(M) + w_c(\rho)$  and  $d(M/\rho) = d(M) + w_d(\rho)$ . Also, it is easy to see that for any closed subset  $R$ ,

$$c(M_R) = c(M_0) + \sum_{\rho \in R} w_c(\rho) \quad \text{and} \quad d(M_R) = d(M_0) + \sum_{\rho \in R} w_d(\rho).$$

Hence, the minimum egalitarian stable marriage problem (the sex-equal stable marriage problem, respectively) is equivalent to the problem of finding a closed subset  $R$  such that  $c(M_0) + \sum_{\rho \in R} w_c(\rho)$  ( $|d(M_0) + \sum_{\rho \in R} w_d(\rho)|$ , respectively) is minimum. For example, the algorithm for finding a minimum egalitarian stable matching in [Irving et al. 1987] efficiently finds such  $R$  by exploiting network flow.

### 3. THE SEX-EQUAL STABLE MARRIAGE PROBLEM

Recall that  $M_0$  is the man-optimal stable matching and  $M_z$  is the woman-optimal stable matching. Note that any stable matching  $M$  satisfies  $d(M_0) \leq d(M) \leq d(M_z)$ . Thus, the sex-equal stable marriage problem is trivial if  $d(M_0) \geq 0$  or  $d(M_z) \leq 0$ , namely, if  $d(M_0) \geq 0$ ,  $M_0$  is optimal, while if  $d(M_z) \leq 0$ ,  $M_z$  is optimal. Therefore, we consider the case where  $d(M_0) < 0 < d(M_z)$ . Recall that  $\Delta = \min\{|d(M_0)|, |d(M_z)|\}$ . In the following, we assume without loss of generality that  $|d(M_0)| \leq |d(M_z)|$ , since otherwise we can exchange the roles of men and women. Hence,  $\Delta = \min\{|d(M_0)|, |d(M_z)|\} = |d(M_0)|$ .

We first briefly give the underlying idea of our algorithm presented in this section. Recall that, for a given instance  $I$  and  $\epsilon$ , we are to find a stable matching  $M$  such that  $-\epsilon\Delta \leq d(M) \leq \epsilon\Delta$  if any. As an easy case, assume that all rotations  $\rho$  of  $I$  satisfy  $w_d(\rho) \leq 2\epsilon\Delta$ . Now, we construct the rotation poset  $(\Pi, \prec)$  of  $I$ , and starting from  $M_0$ , we eliminate rotations in an order of any linear extension of  $\prec$ . Recall that by eliminating a rotation, the sex-equality cost increases, but by at most  $2\epsilon\Delta$  by assumption. Note that  $d(M_0) < 0 < d(M_z)$ , and recall that if we eliminate all rotations from  $M_0$ , we eventually reach  $M_z$ . Then, in this sequence, we certainly meet a desirable stable matching at some point.

However, this procedure fails if there is a rotation with large sex-equality cost: If we eliminate such a rotation, then we may “jump” from  $M$  to  $M'$  such that  $d(M) < -\epsilon\Delta$  and  $d(M') > \epsilon\Delta$  even if there is a feasible solution. To resolve this problem, we will try all combinations of selecting such “large” rotations, and treat “small” rotations in the above manner. To evaluate the time complexity, we show that the number of large rotations used in a feasible solution is limited.

Before giving a description of our algorithm, we define our notation. Let  $R$  be any (not necessarily closed) subset of a poset  $(\Pi, \prec)$ . Then define  $R_{\min} = R \cup \{\rho \mid \text{there exists a } \rho' \in R \text{ such that } \rho \prec \rho'\}$  and  $R_{\max} = R \cup \{\rho \mid \text{there exists a } \rho' \in R \text{ such that } \rho' \prec \rho\}$ . See Fig. 2 for examples. Fig. 2 (a) represents a Hasse diagram of a poset, where the black circles are elements of  $R$ .  $R_{\min}$  and  $R_{\max}$  are depicted in Fig. 2 (b) and Fig. 2 (c), respectively. Intuitively speaking, when constructing a closed subset  $A$ , if we decide to include all elements of  $R$  to  $A$ , then  $R_{\min}$  is the set of elements that must be included in  $A$ . Similarly, if we decide to include no elements of  $R$ , then  $R_{\max}$  is the set of elements that must not be included in  $A$ .

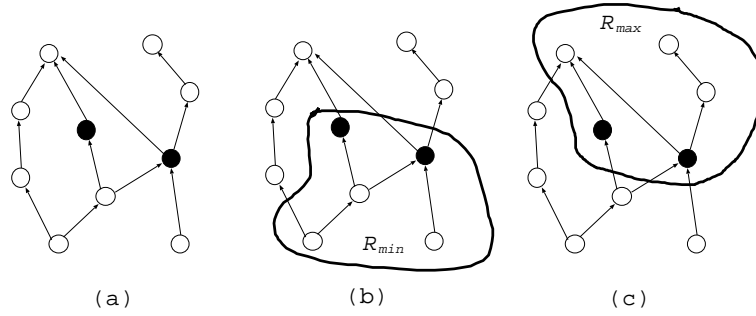


Fig. 2. Examples of  $R_{\min}$  and  $R_{\max}$ .

---

#### Algorithm 1

1. Construct the rotation poset  $(\Pi, \prec)$ .
  2. Let  $R^L$  be the set of rotations  $\rho$  such that  $w_d(\rho) > 2\epsilon\Delta$ , and  $R^S$  be  $\Pi \setminus R^L$ .
  3. For each set  $R$  in  $2^{R^L}$  such that  $|R| \leq \frac{1+\epsilon}{2\epsilon}$ , do,
    - (a) If  $-\epsilon\Delta \leq d(R_{\min}) \leq \epsilon\Delta$ , then output  $M_{R_{\min}}$ .
    - (b) Fix an arbitrary order  $\rho_1, \rho_2, \dots, \rho_k$  which is consistent with  $\prec$ , where  $\{\rho_1, \rho_2, \dots, \rho_k\} = R^S \setminus (R_{\min} \cup (R^L \setminus R)_{\max})$ .  
(Note that  $R^S \setminus (R_{\min} \cup (R^L \setminus R)_{\max})$  could be empty.)
    - (c) For  $i = 1$  to  $k$   
 if  $-\epsilon\Delta \leq d(R_{\min} \cup \{\rho_1, \rho_2, \dots, \rho_i\}) \leq \epsilon\Delta$ , then output  $M_{R_{\min} \cup \{\rho_1, \rho_2, \dots, \rho_i\}}$   
 and halt.
  4. Output “None” and halt.
- 

**THEOREM 3.1.** *There is an algorithm for NSE whose running time is  $O(n^{3+\frac{1}{\epsilon}})$ .*

**PROOF. Correctness Proof.** Clearly, if there is no stable matching  $M$  such that  $-\epsilon\Delta \leq d(M) \leq \epsilon\Delta$ , then the algorithm answers “None.” Otherwise, suppose that there is a stable matching  $M_X$  such that  $-\epsilon\Delta \leq d(M_X) \leq \epsilon\Delta$ , where  $X$  is the set of rotations corresponding to  $M_X$ . Let  $X^L = X \cap R^L$  and  $X^S = X \cap R^S$ . Then,  $d(X^L) \leq d(X) = d(M_X) \leq \epsilon\Delta$ . Note that  $d(X^L) = d(M_0) + \sum_{\rho \in X^L} w_d(\rho)$ . Also,  $w_d(\rho) > 2\epsilon\Delta$  for any rotation  $\rho \in X^L$ . It then follows that  $|X^L| < \frac{d(X^L) - d(M_0)}{2\epsilon\Delta} \leq$



$\frac{|d(M_0)| + \epsilon\Delta}{2\epsilon\Delta} = \frac{1+\epsilon}{2\epsilon}$ . Therefore, Algorithm 1 selects  $X^L$  at Step 3 as  $R$ , and we consider this particular execution of Step 3.

First, note that  $d((X^L)_{\min}) \leq \epsilon\Delta$ , since otherwise  $d(X) \geq d((X^L)_{\min}) > \epsilon\Delta$ , a contradiction. If  $-\epsilon\Delta \leq d((X^L)_{\min}) \leq \epsilon\Delta$ , then Algorithm 1 outputs  $M_{(X^L)_{\min}}$  at Step 3(a). Therefore, suppose that  $d((X^L)_{\min}) < -\epsilon\Delta$ . Note that  $X^L \subseteq (X^L)_{\min}$  by definition, and  $X^S \subseteq (X^L)_{\min} \cup \{\rho_1, \rho_2, \dots, \rho_k\}$ . Since  $X = X^L \cup X^S$ ,  $X \subseteq (X^L)_{\min} \cup \{\rho_1, \rho_2, \dots, \rho_k\}$ . Therefore,  $d((X^L)_{\min} \cup \{\rho_1, \rho_2, \dots, \rho_k\}) \geq d(X) \geq -\epsilon\Delta$ . Note also that any rotation  $\rho_i$  ( $1 \leq i \leq k$ ) satisfies  $w_d(\rho_i) \leq 2\epsilon\Delta$ . Hence there must be  $j$  ( $1 \leq j \leq k$ ) such that  $-\epsilon\Delta \leq d((X^L)_{\min} \cup \{\rho_1, \rho_2, \dots, \rho_j\}) \leq \epsilon\Delta$ .

Finally, we show that  $R_{\min} \cup \{\rho_1, \rho_2, \dots, \rho_i\}$  at Step 3(c) is closed. For this, it suffices to show that  $R_{\min} \cup \{\rho_1, \rho_2, \dots, \rho_k\}$  is closed. Suppose that it is not closed. Then, since  $R_{\min}$  is closed by definition, there are rotations  $\rho \notin R_{\min} \cup \{\rho_1, \rho_2, \dots, \rho_k\}$  and  $\rho_j$  ( $1 \leq j \leq k$ ) such that  $\rho \prec \rho_j$ . The fact that  $\rho \notin R_{\min} \cup \{\rho_1, \rho_2, \dots, \rho_k\}$  implies that  $\rho \in (R^L \setminus R)_{\max}$ , but then  $\rho_j \in (R^L \setminus R)_{\max}$  by the definition of  $(R^L \setminus R)_{\max}$ . This contradicts the fact that  $\rho_j \in R^S \setminus (R_{\min} \cup (R^L \setminus R)_{\max})$ .

**Time Complexity.** Steps 1 and 2 can be performed in  $O(n^2)$ . Inside the loop of Step 3 can be performed in  $O(n^2)$  since the number of rotations is at most  $O(n^2)$ . Clearly, Step 4 can be performed in constant time.

We consider the number of repetitions of Step 3, i.e., the number of  $R$  satisfying the condition at Step 3. Let this number be  $t$ . Recall that the number of rotations is at most  $n^2$  as mentioned in Sec. 2. Therefore,  $|R^L| \leq n^2$ . Since  $|R| \leq \frac{1+\epsilon}{2\epsilon}$ ,

$$t = \sum_{k=1}^{\lfloor \frac{1+\epsilon}{2\epsilon} \rfloor} \binom{n^2}{k} \leq \sum_{k=1}^{\lfloor \frac{1+\epsilon}{2\epsilon} \rfloor} \frac{(n^2)^{\lfloor \frac{1+\epsilon}{2\epsilon} \rfloor}}{k!} = O(n^{\frac{1+\epsilon}{\epsilon}}).$$

Hence the time complexity of Algorithm 1 is  $O(n^2) \cdot O(n^{\frac{1+\epsilon}{\epsilon}}) = O(n^{3+\frac{1}{\epsilon}})$ .  $\square$

**Remark 1.** The Partially Ordered Knapsack problem (POK) [Johnson and Niemi 1983; Kolliopoulos and Steiner 2007] is the Knapsack problem with the following additional constraints. There is a precedence relation between items, and items must be selected according to this relation. Once the rotation poset is constructed, the sex-equal stable marriage problem is essentially the same as a special case of POK, denoted UPOK, where every item has the same profit and weight. Very recently, Bonsma [Bonsma 2007], independently of us, proposed a PTAS for a special case of UPOK using the same idea as Algorithm 1.

**Remark 2.** We can improve Algorithm 1 when  $|d(M_0)|$  and  $|d(M_z)|$  are close, more precisely, when they differ by at most a  $\log n$  factor. Let  $\Delta'$  be  $\frac{1}{\log n} \max\{|d(M_0)|, |d(M_z)|\}$ . We can find a stable matching  $M$  which satisfies  $-\epsilon\Delta' \leq d(M) \leq \epsilon\Delta'$  or report that none exists in polynomial time by using a modified version of Algorithm 1 (Algorithm 1'). We modify Algorithm 1 so that it uses  $\Delta'$  instead of  $\Delta$  and executes Step 3 for all subsets of  $2^{R^L}$ . Note that, from the discussion in Sec. 2,

$$d(M_z) = d(M_0) + \sum_{\rho \in \Pi} w_d(\rho).$$

Hence  $|R^L| < \frac{d(M_z) - d(M_0)}{2\epsilon\Delta'} \leq \frac{2 \max\{|d(M_0)|, |d(M_z)|\}}{2\epsilon\Delta'} = \frac{\log n}{\epsilon}$ . Therefore the number



of repetitions of Step 3 is at most  $2^{|R^L|} = n^{1/\epsilon}$  and hence the time complexity of Algorithm 1' is  $O(n^{2+\frac{1}{\epsilon}})$ .

**Remark 3.** There are several goodness measures of an approximation algorithm  $A$  for a minimization problem. The usual measure is the *approximation ratio* of  $A$ , which is defined as  $\max\{A(x)/opt(x)\}$  over all instances  $x$ , where  $opt(x)$  and  $A(x)$  are the costs of the optimal and the algorithm's solutions, respectively. However, this measure cannot be used for the sex-equal stable marriage problem, because  $opt(x)$  can be zero. For such a case, there is another measure: the *relative accuracy* [Charikar and Wirth 2004; Nesterov 1998], which is defined as  $\max\{(\max(x) - opt(x))/(\max(x) - A(x))\}$  over all instances  $x$ , where  $opt(x)$ ,  $\max(x)$ , and  $A(x)$  are the costs of the optimal solution, the worst solution, and the algorithm's solution, respectively. By using Algorithm 1' in Remark 2, we can construct an algorithm  $T$  which achieves the relative accuracy  $1 + \epsilon/\log n$  for an arbitrary constant  $\epsilon > 0$ .

To see this, let  $M_{opt}$  be an optimal solution for the sex-equal stable marriage problem. Recall that we are considering the case where  $d(M_0) < 0 < d(M_z)$ . If  $|d(M_{opt})| > D/2$ , where  $D = \max\{|d(M_0)|, |d(M_z)|\}$ ,  $M_{opt}$  can be obtained in polynomial time in the following way: Let  $M_a$  be the stable matching such that no other stable matching  $M$  satisfies  $d(M_a) < d(M) < -D/2$  and let  $M_b$  be the stable matching such that no other stable matching  $M$  satisfies  $D/2 < d(M) < d(M_b)$ . Then,  $M_{opt}$  is either  $M_a$  or  $M_b$ . Since  $d(M_b) - d(M_a) > D$ , there exists a rotation  $\rho_H$  such that  $w_d(\rho_H) > D$  (otherwise there must be a stable matching  $M_c$  such that  $d(M_a) < d(M_c) < d(M_b)$ ). Also, this  $\rho_H$  is unique because

$$\sum_{\rho \in \Pi} w_d(\rho) = d(M_z) - d(M_0) \leq 2D.$$

It is easy to see that the maximal closed subset which does not contain  $\rho_H$  corresponds to  $M_a$  and that the minimal closed subset which contains  $\rho_H$  corresponds to  $M_b$ . Therefore,  $M_a$  and  $M_b$  can be obtained in polynomial time. Finally, assume that  $|d(M_{opt})| \leq D/2$ . For each  $i$  such that  $i = 1, 2, \dots, \lceil \frac{\log n}{\epsilon} \rceil$ , we can find a stable matching with sex-equality cost between  $-\frac{\epsilon}{2 \log n} Di$  and  $\frac{\epsilon}{2 \log n} Di$  if any by using Algorithm 1' (by adjusting  $\Delta'$  appropriately), and output the best one. Then, it is easy to see that an output matching  $M$  satisfies  $|d(M)| - |d(M_{opt})| \leq \frac{\epsilon}{2 \log n} D$ . Now, the relative accuracy is

$$\frac{\max(x) - opt(x)}{\max(x) - T(x)} = 1 + \frac{T(x) - opt(x)}{\max(x) - T(x)} \leq 1 + \frac{(\epsilon/2 \log n)D}{D/2} = 1 + \frac{\epsilon}{\log n}.$$

#### 4. THE MINIMUM EGALITARIAN SEX-EQUAL STABLE MARRIAGE PROBLEM

In NSE, we are asked to find a stable matching whose sex-equality cost is in some range close to 0. However, if there are several stable matchings satisfying the condition, there might be good ones and bad ones. In fact, there is an instance  $I$  that has two stable matchings  $M$  and  $M'$  whose sex-equality costs are the same (0), but whose egalitarian costs are significantly different. Instance  $I$  is constructed with the following steps. First consider the following instance  $I_1$  consisting of  $2n$  men and  $2n$  women:

$m_1$ :	$w_1 w_{n+1} w_{n+2} \cdots w_{2n} w_2$	$w_1$ :	$m_n m_1$
$m_2$ :	$w_2 w_{n+1} w_{n+2} \cdots w_{2n} w_3$	$w_2$ :	$m_1 m_2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$m_{n-1}$ :	$w_{n-1} w_{n+1} w_{n+2} \cdots w_{2n} w_n$	$w_{n-1}$ :	$m_{n-2} m_{n-1}$
$m_n$ :	$w_n w_{n+1} w_{n+2} \cdots w_{2n} w_1$	$w_n$ :	$m_{n-1} m_n$
$m_{n+1}$ :	$w_{n+1}$	$w_{n+1}$ :	$m_{n+1} m_1 \cdots m_n$
$m_{n+2}$ :	$w_{n+2}$	$w_{n+2}$ :	$m_{n+2} m_1 \cdots m_n$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$m_{2n-1}$ :	$w_{2n-1}$	$w_{2n-1}$ :	$m_{2n-1} m_1 \cdots m_n$
$m_{2n}$ :	$w_{2n}$	$w_{2n}$ :	$m_{2n} m_1 \cdots m_n$

If a preference list is not complete, then add any missing persons at the tail of the list in an arbitrary order. Instance  $I_1$  has two stable matchings:  $M_1 = \{(m_1, w_1), (m_2, w_2), \dots, (m_{n-1}, w_{n-1}), (m_n, w_n), (m_{n+1}, w_{n+1}), (m_{n+2}, w_{n+2}), \dots, (m_{2n-1}, w_{2n-1}), (m_{2n}, w_{2n})\}$  and  $M_2 = \{(m_1, w_2), (m_2, w_3), \dots, (m_{n-1}, w_n), (m_n, w_1), (m_{n+1}, w_{n+1}), (m_{n+2}, w_{n+2}), \dots, (m_{2n-1}, w_{2n-1}), (m_{2n}, w_{2n})\}$ . Note that  $c(M_1) = 5n$ ,  $d(M_1) = -n$ ,  $c(M_2) = n^2 + 5n$ , and  $d(M_2) = n^2 + n$ . Let  $I_2$  be the instance obtained from  $I_1$  by exchanging men and women.

Let  $I$  be the instance obtained by putting  $I_1$  and  $I_2$  together and padding the missing persons at the tail of the preference lists to make them complete. In more detail, the set of men of  $I$  is the union of the sets of men in  $I_1$  and  $I_2$ , and the set of women of  $I$  is similarly defined. The preference list of a man  $m$  in  $I$  who came from  $I_1$  is constructed by adding the women in  $I_2$  in an arbitrary order to the tail of the list of  $m$  in  $I_1$ . The preference lists of the other people are constructed similarly. Then,  $I$  has four stable matchings  $M_3$ ,  $M_4$ ,  $M_5$ , and  $M_6$ , whose egalitarian costs and sex-equality costs are given in Table I. We see that  $d(M_3) = d(M_6) = 0$ , while  $c(M_3)$  is small but  $c(M_6)$  is large.

	$M_3$	$M_4$	$M_5$	$M_6$
$c(M)$	$10n$	$n^2 + 10n$	$n^2 + 10n$	$2n^2 + 10n$
$d(M)$	$0$	$-n^2 - 2n$	$n^2 + 2n$	$0$

Table I. The egalitarian costs and the sex-equality costs

This motivates us to consider the following problem, *MinESE (the Minimum Egalitarian Sex-Equal stable marriage problem)*: Given an instance  $I$  and a constant  $\epsilon$  such that  $0 < \epsilon < 1$ , find a stable matching  $M$  with minimum  $c(M)$ , under the condition that  $|d(M)| \leq \epsilon\Delta$ , (or answer “None” if no such solution exists). First, in Sec. 4.1, we show that MinESE is strongly NP-hard. Then, in Sec. 4.2, we give an approximation algorithm for MinESE.

#### 4.1 Strongly NP-hardness of MinESE

It turned out that there is a polynomial-time algorithm for obtaining a stable matching  $M$  such that (a)  $-\epsilon\Delta \leq d(M) \leq \epsilon\Delta$  or (b)  $c(M)$  is minimum. Interestingly, it is hard to obtain  $M$  satisfying (a) and (b).

**THEOREM 4.1.** *MinESE is strongly NP-hard.*

PROOF. We will prove the theorem by a reduction from the clique problem. First, we give the definition of the clique problem.

---

**The clique problem**

**Input:** A graph  $G(V, E)$  and a positive integer  $k$ .

**Output:** “Yes,” if  $G(V, E)$  has a clique of size  $k$ .  
“No,” otherwise.

---

This problem is NP-complete even if we restrict the problem so that  $G$  is  $d$ -regular with  $d > 2k^2$ . This can be seen, for example, in the following way: Given  $G = (V, E)$ , define  $d$  as the smallest even number that satisfies  $d \geq \max\{\Delta(G), 2k^2 + 1\}$ , where  $\Delta(G)$  is the maximum degree of  $G$ . Let  $V = \{v_1, v_2, \dots, v_n\}$ , and for each  $v_i$ , let  $d(v_i)$  be its degree in  $G$ . For each  $v_i$ , define  $\tilde{d}(v_i) = d - d(v_i)$ , and let  $\tilde{d}(G) = \sum_i \tilde{d}(v_i) = dn - \sum_i d(v_i)$ . Note that  $\tilde{d}(G)$  is even since both  $dn$  and  $\sum_i d(v_i)$  are even. Let  $s = \lceil \frac{\tilde{d}(G)}{2d} \rceil$ , and for  $1 \leq i \leq s$ , construct a complete bipartite graph  $G_i = (A_i, B_i, E_i)$  where  $A_i = \{a_1^i, a_2^i, \dots, a_d^i\}$  and  $B_i = \{b_1^i, b_2^i, \dots, b_d^i\}$  are vertex sets ( $|A_i| = |B_i| = d$ ). Note that each  $G_i$  is  $d$ -regular. Next, for each  $G_i$  ( $1 \leq i \leq s-1$ ), remove the edges  $(a_j^i, b_j^i)$  for  $1 \leq j \leq d$ , and for  $G_s$ , remove the edges  $(a_j^s, b_j^s)$  for  $1 \leq j \leq \frac{\tilde{d}(G) - 2d(s-1)}{2}$ . Now there are  $\tilde{d}(G)$  vertices of degree  $d-1$  in  $G_1, G_2, \dots, G_s$ . Connect these  $\tilde{d}(G)$  vertices with vertices in the original graph  $G$  in an arbitrary way so that all of the vertices have degree  $d$ . It is easy to see that this procedure can be done in polynomial time and does not affect the existence of the  $k$ -clique.

In the following, we assume that the given graph has the above property, i.e.,  $d$ -regular with  $d > 2k^2$ . First, we construct a poset from the clique problem. Given a graph  $G = (V, E)$  and an integer  $k$ , we construct a poset  $(\Pi, <)$  in a similar manner as the construction used in [Johnson and Niemi 1983]. Let  $\Pi$  be  $V \cup E \cup \{\rho_+\}$ , where  $\rho_+$  is an additional element. Define the precedence relation  $<$  as follows:  $\rho_+ < v$  for all  $v \in V$ , and  $v < e$  if and only if  $v \in V$  is incident to  $e \in E$  in  $G(V, E)$ . Then,  $\rho_+$  has outdegree  $|V|$ , each element  $v \in V$  has outdegree  $d$  and indegree 1, and each element  $e \in E$  has indegree 2. Note that, to construct a non-empty closed subset, we need to choose  $\rho_+$ , and if we select  $k$  elements from  $V$ , we can take at most  $k(k-1)/2$  elements from  $E$ .

Next, we construct a MinESE instance  $I'$  from the poset  $(\Pi, <)$  obtained by this construction. We first construct  $I$  from  $(\Pi, <)$  using the same construction as [Gusfield et al. 1987; Kato 1993]. In the following, we give a brief explanation how the instance  $I$  is constructed:

- Construct a poset  $(\Pi', <')$  from  $(\Pi, <)$  by adding a new element  $s$  that precedes all elements in  $\Pi$ , and another new element  $t$  that succeeds all elements in  $\Pi$ .
- Rename elements of  $\Pi'$  as  $\rho_1, \rho_2, \dots, \rho_h$  ( $h = |\Pi'|$ ), so that the order  $\rho_1, \rho_2, \dots, \rho_h$  is a linear extension of  $<'$ ; in other words, for all  $i$  and  $j$ , if  $\rho_i < \rho_j$  then  $i < j$ .
- Let  $\{e_1, e_2, \dots, e_F\}$  be the set of edges of the Hasse diagram  $H(\Pi')$  of  $(\Pi', <')$ . Associated with each edge  $e_i$ , create a man  $m_i$  and a woman  $w_i$ . There will be  $F$  men and  $F$  women in total.

- For each  $i$  ( $1 \leq i \leq F$ ), include  $w_i$  in  $m_i$ 's list, and similarly include  $m_i$  in  $w_i$ 's list.
- For  $i = 1$  to  $h$ , do the following.
  - Let  $r_i$  be the degree of  $\rho_i$  in  $(\Pi', \prec')$ .
  - Let  $(m_0^i, m_1^i, \dots, m_{r_i-1}^i)$  be an arbitrary ordering of men corresponding to edges incident to  $\rho_i$  in  $H(\Pi')$ .
  - Let  $(w_0^i, w_1^i, \dots, w_{r_i-1}^i)$  be the ordered set of women such that  $w_j^i$  is the last woman on  $m_j^i$ 's current preference list.
  - For  $j = 0$  to  $r_i - 1$ , place  $w_{j+1}^i$  at the end of  $m_j^i$ 's list, and place  $m_j^i$  at the head of  $w_{j+1}^i$ 's list, where  $j+1$  is taken modulo  $r_i$ .
- Add any missing entries at the end of each person's list in an arbitrary order (to make the lists complete).

The obtained stable marriage instance  $I$  has the following properties:

- (a) The rotation poset of  $I$  is exactly  $(V \cup E \cup \{\rho_+\}, \prec)$ .
- (b) The rotation  $\rho_+$  involves  $|V| + 1$  men and  $|V| + 1$  women.
- (c) Let  $\rho_v$  be a rotation corresponding to  $v \in V$ . Then,  $\rho_v$  involves  $d + 1$  men and  $d + 1$  women.
- (d) Let  $\rho_e$  be a rotation corresponding to  $e \in E$ . Then,  $\rho_e$  involves three men and three women.
- (e) For each rotation  $\rho = \{(m_0, w_0), \dots, (m_{r-1}, w_{r-1})\}$  of  $I$  and for each  $i$ ,  $w_{i+1}$  is next to  $w_i$  in  $m_i$ 's list and  $m_i$  is next to  $m_{i-1}$  in  $w_i$ 's list, where  $i+1$  and  $i-1$  are taken modulo  $r$ .

It should be noted that  $w_c(\rho) = 0$  for all rotations  $\rho$  in  $I$  by property (e). Then, we modify  $I$  and construct  $I'$  so that the following holds in  $I'$ :

- (1)  $w_c(\rho_+) = 0$  and  $w_d(\rho_+) = ((1 + \epsilon)D - 2B)/(1 - \epsilon)$ , where  $B = 2dk + 4k^2 - 2k$  and  $D = (2d + 2)|V| + 8|E|$ . Here,  $\epsilon$  is any constant such that  $0 < \epsilon < 1$ , which we can determine as an input of MinESE.
- (2) For any rotation  $\rho_v$  corresponding to  $v \in V$ ,  $w_c(\rho_v) = 0$  and  $w_d(\rho_v) = 2d + 2$ .
- (3) For any rotation  $\rho_e$  corresponding to  $e \in E$ ,  $w_c(\rho_e) = -2$  and  $w_d(\rho_e) = 8$ .
- (4) The man-optimal stable matching  $M_0$  has the sex-equality cost  $-(D - B)/(1 - \epsilon)$ , so,  $d(M_0) = -(D - B)/(1 - \epsilon)$  (note that  $B$  and  $D$  are defined in (1)).

The condition (2) has already been satisfied. The two conditions (1) and (3) can be satisfied by padding “dummy” persons in the preference lists as in [Kato 1993]: For the rotation  $\rho_+ = (m_1, w_1), \dots, (m_{|V|+1}, w_{|V|+1})$ , we pad  $((1 + \epsilon)D - 2B)/2(1 - \epsilon) - (|V| + 1)$  (dummy) women between  $w_1$  and  $w_2$  in  $m_1$ 's list in  $I'$  and  $((1 + \epsilon)D - 2B)/2(1 - \epsilon) - (|V| + 1)$  (dummy) men between  $m_1$  and  $m_2$  in  $w_1$ 's list in  $I'$ . For each rotation  $\rho_e = (m_1, w_1), (m_2, w_2), (m_3, w_3)$  corresponding to  $e \in E$ , we pad 2 (dummy) men between  $m_1$  and  $m_2$  in  $w_1$ 's list, with the result that  $w_c(\rho_e) = -2$  and  $w_d(\rho_e) = 8$  in  $I'$ . The last condition (4) can be satisfied by padding dummy men (or women) at the top of the women's (or men's) lists. (For example, if we want to decrease the value of  $d(M_0)$  without violating the conditions (1)–(3), we pad (dummy) men to the top of the women's lists.) The resulting instance of MinESE

is  $(I', \epsilon)$ . Note that these constructions can be done in polynomial time. It is easy to check that in the instance  $I'$ ,  $d(M_z) = d(M_0) + w_d(\rho_+) + ((2d+2)|V| + 8|E|) = -(D-B)/(1-\epsilon) + ((1+\epsilon)D - 2B)/(1-\epsilon) + D = (D-B)/(1-\epsilon)$ . Hence,  $|d(M_0)| = |d(M_z)|$  and

$$\begin{aligned}\epsilon\Delta &= \epsilon \min\{|d(M_0)|, |d(M_z)|\} \\ &= \epsilon |d(M_0)| \\ &= \epsilon(D-B)/(1-\epsilon) \\ &= d(M_0) + w_d(\rho_+) + B.\end{aligned}\tag{1}$$

We will show that  $G$  has a  $k$ -clique if and only if there is a stable matching  $M$  in  $I'$  such that  $-\epsilon\Delta \leq d(M) \leq \epsilon\Delta$  and  $c(M) \leq c(M_0) - k(k-1)$ . If this is true, we can show that MinESE is strongly NP-hard as follows: Given an instance  $G$  of the clique problem, we construct a MinESE instance  $I'$  and  $\epsilon$  by the above reduction. Then, we find an optimal solution  $M$  and the man-optimal stable matching  $M_0$ . Finally, we compare  $c(M)$  and  $c(M_0) - k(k-1)$ : If  $c(M) \leq c(M_0) - k(k-1)$ , then the answer to the clique problem is “Yes,” otherwise, “No.”

We first show the “only if” part. Suppose that  $G$  has a  $k$ -clique  $C$ . In the rotation poset of  $I'$ , let  $R$  be the set of rotations corresponding to  $k$  vertices and  $k(k-1)/2$  edges of  $C$ . Then,  $R' = R \cup \{\rho_+\}$  is a closed subset. Then, it is easy to see that the corresponding stable matching  $M_{R'}$  is a required solution:  $c(M_{R'}) = c(M_0) + 0 + 0 \cdot (k-1) - 2k(k-1)/2 = c(M_0) - k(k-1)$ . Also,  $d(M_{R'}) = d(M_0) + w_d(\rho_+) + k(2d+2) + 8k(k-1)/2 = d(M_0) + w_d(\rho_+) + B = \epsilon\Delta$  (the last equality is from Equation (1)).

We now show the “if” part. Let  $M$  be any stable matching of  $I'$  such that  $-\epsilon\Delta \leq d(M) \leq \epsilon\Delta$  and  $c(M) \leq c(M_0) - k(k-1)$ , and  $R$  be the set of rotations of  $I'$  corresponding to  $M$ . First, note that  $R$  contains  $\rho_+$ , since otherwise,  $R = \emptyset$  and so  $d(M) = d(M_0) = -\Delta$ , which contradicts the assumption that  $d(M) \geq -\epsilon\Delta$ . Let  $v_M$  and  $e_M$  be the numbers of the rotations in  $R$  which correspond to elements in  $V$  and  $E$ , respectively. Then,  $d(R) = d(M) = d(M_0) + w_d(\rho_+) + (2d+2)v_M + 8e_M$ . If  $v_M \geq k+1$ ,  $d(M) > d(M_0) + w_d(\rho_+) + B$  since  $d > 2k^2$ . Hence,  $d(M) > \epsilon\Delta$ , again contradicting the assumption. Therefore,  $v_M$  must satisfy  $0 \leq v_M \leq k$ . Recall that  $e_M \leq v_M(v_M-1)/2$ . Suppose that  $0 \leq e_M < k(k-1)/2$ . Then,  $c(M) = c(M_0) + 0 \cdot v_M - 2e_M > c(M_0) - k(k-1)$ , again contradicting the assumption. Thus,  $e_M = k(k-1)/2$  and  $v_M = k$ . Therefore, the vertices and the edges in  $R$  correspond to a  $k$ -clique of  $G$ .  $\square$

**Remark 4.** Note that the reduction in the NP-hardness proof produces an instance  $(I, \epsilon)$  of MinESE such that  $|d(M_0)| = |d(M_z)|$  in  $I$ , and  $\epsilon$  is any constant such that  $0 < \epsilon < 1$ . Observe that if  $|d(M_0)| = |d(M_z)|$  and  $\epsilon = 1$ , then MinESE is equivalent to the minimum egalitarian stable marriage problem, which can be solved in polynomial time.

#### 4.2 Approximation Algorithms for MinESE

Here, we give a  $(2-(\epsilon-\delta)/(2+3\epsilon))$ -approximation algorithm for MinESE for an arbitrary  $\delta$  such that  $0 < \delta < \epsilon$ . Similarly to Sec. 3, we assume that  $|d(M_0)| \leq |d(M_z)|$ . We begin by stating two simple but important results that link the egalitarian cost

and the sex-equality cost; their proofs are given later. (i) For any stable matching  $M$ ,  $|d(M)| < c(M)$  (Lemma 4.4). (ii) For any stable matching  $M$  and a rotation  $\rho$  exposed in  $M$ , by eliminating  $\rho$  from  $M$ , the cost change with respect to the egalitarian measure is less than the cost change with respect to the sex-equality measure (Lemma 4.5).

To illustrate the basic idea of the algorithm, we first consider a restricted case and show that our algorithm achieves a 2-approximation. For a fixed  $\delta$  such that  $\epsilon > \delta > 0$ , suppose that all of the rotations satisfy  $w_d(\rho) \leq \delta\Delta$ . Given  $I$  and  $\epsilon$ , we first find a minimum egalitarian stable matching  $M_{eg}$ , which can be done in polynomial time. If  $-\epsilon\Delta \leq d(M_{eg}) \leq \epsilon\Delta$ , then we are done since  $M_{eg}$  is an optimal solution for MinESE. If  $d(M_{eg}) < -\epsilon\Delta$ , then we eliminate rotations one by one as Algorithm 1 does until the sex-equality cost first becomes  $-\epsilon\Delta$  or larger. If  $d(M_{eg}) > \epsilon\Delta$ , then we “add” rotations one by one until the sex-equality cost first becomes  $\epsilon\Delta$  or smaller. Here, “adding a rotation” means the reverse operation of eliminating a rotation. If we do not reach a feasible solution by this procedure, then we can conclude that there is no feasible solution, by a similar argument as in Sec. 3. If we find a stable matching  $M$  such that  $-\epsilon\Delta \leq d(M) \leq \epsilon\Delta$ , then we can show that this is a 2-approximation, namely,  $c(M) \leq 2c(M_{eg})$  using (i) and (ii) above (note that the optimal cost is at least  $c(M_{eg})$ ): Suppose, for example, that  $d(M_{eg}) < -\epsilon\Delta$  (see Fig. 3). Then, by (ii),  $c(M) - c(M_{eg}) < d(M) - d(M_{eg})$ , and by (i),  $|d(M_{eg})| < c(M_{eg})$ . Also, since the costs of rotations are at most  $\delta\Delta$ , and since  $M$  is the first feasible solution found by this procedure,  $d(M) \leq -(\epsilon - \delta)\Delta < 0$ . Putting these together, we have that

$$\begin{aligned} c(M) &< c(M_{eg}) + d(M) - d(M_{eg}) \\ &< c(M_{eg}) - d(M_{eg}) && (\text{since } d(M) < 0) \\ &= c(M_{eg}) + |d(M_{eg})| && (\text{since } d(M_{eg}) < 0) \\ &< 2c(M_{eg}) && (\text{since } |d(M_{eg})| < c(M_{eg})). \end{aligned}$$

Hence  $c(M)/c(M_{eg}) < 2$ .

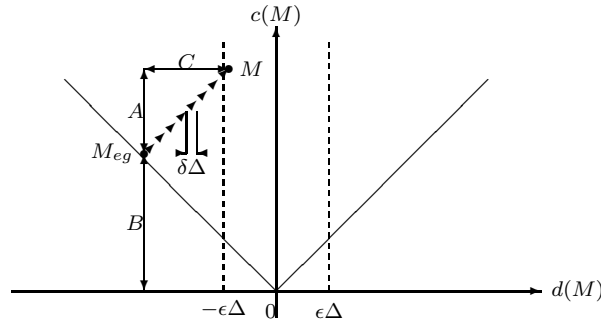


Fig. 3.  $C < B$  by (i) and  $A < C$  by (ii). Hence  $A + B < C + B < 2B$ .

However, we may have rotations of large costs. Then we take a similar approach as in Sec. 3: Let  $R^L$  be the set of such large rotations. Then, for any partition  $R_1$  and  $R_2$  of  $R^L$  ( $R_1 \cup R_2 = R^L$  and  $R_1 \cap R_2 = \emptyset$ ), we want to obtain a minimum

egalitarian stable matching whose corresponding closed subset  $A$  contains all of the rotations in  $R_1$  but none in  $R_2$ . For this purpose, we need to solve the following problem:

**Instance.** An instance  $I$  of the stable marriage problem, and its disjoint subsets of rotations  $R_1$  and  $R_2$ , both of which are subsets of  $R^L$ .

**Feasible solution.** A closed subset  $A$  such that  $A \supseteq R_1$  and  $A \cap R_2 = \emptyset$ .

**Optimization criteria.** Minimize the egalitarian cost of the stable matching  $M_A$  corresponding to  $A$ .

For this problem, we can use the same algorithm for the minimum egalitarian stable marriage problem in [Gusfield and Irving 1989]. We denote this procedure by  $\text{minEgalitarian}(R_1, R_2)$ . First, we review the following proposition:

**PROPOSITION 4.2.** [Feder 1992; 1994] *Given a poset  $(\Pi, \prec)$ , there is an  $O(n^3)$ -time algorithm which finds a minimum-weight closed subset of  $(\Pi, \prec)$  with respect to the egalitarian cost.*

Our procedure  $\text{minEgalitarian}(R_1, R_2)$  is as follows: Without loss of generality, assume that there are no elements such that  $r_2 \prec r_1$  ( $r_1 \in R_1$  and  $r_2 \in R_2$ ) since there exists no solution in such a case. Construct the poset  $(\Pi', \prec)$  by removing all of the rotations in  $(R_1)_{\min}$  and  $(R_2)_{\max}$  from  $(\Pi, \prec)$  (recall the definitions of  $R_{\min}$  and  $R_{\max}$  given before Algorithm 1), and let  $R'$  be the subset obtained by applying Proposition 4.2 to  $(\Pi', \prec)$ . Then, it is easy to see that  $(R_1)_{\min} \cup R'$  is an optimal solution for  $\text{minEgalitarian}(R_1, R_2)$ . Now, we are ready to give the algorithm for MinESE.

---

### Algorithm 2

1. Construct the rotation poset  $(\Pi, \prec)$ .
  2. Let  $M_{\text{best}} = \text{NULL}$ .
  3. Let  $R^L$  be the set of rotations  $\rho$  such that  $w_d(\rho) > \delta\Delta$ , and  $R^S$  be  $\Pi \setminus R^L$ .
  4. For each set  $R$  in  $2^{R^L}$  such that  $|R| \leq \frac{1+\epsilon}{\delta}$ , do,
    - (a) Let  $A = \text{minEgalitarian}(R, R^L \setminus R)$ .
    - (b) If  $d(A) < -\epsilon\Delta$ , do the following.
 

Fix an arbitrary order  $\rho_1, \rho_2, \dots, \rho_k$  which is consistent with  $\prec$ , where  $\{\rho_1, \rho_2, \dots, \rho_k\} = R^S \setminus (A \cup (R^L \setminus R)_{\max})$ .  
(Note that  $R^S \setminus (A \cup (R^L \setminus R)_{\max})$  could be empty.)

For  $i = 1$  to  $k$ , if  $-\epsilon\Delta \leq d(A \cup \{\rho_1, \rho_2, \dots, \rho_i\}) \leq \epsilon\Delta$ , then let  $A = A \cup \{\rho_1, \rho_2, \dots, \rho_i\}$ .

Else if  $d(A) > \epsilon\Delta$ , do the following.

Fix an arbitrary order  $\rho_1, \rho_2, \dots, \rho_k$  which is consistent with  $\prec$ , where  $\{\rho_1, \rho_2, \dots, \rho_k\} = (A \cap R^S) \setminus R_{\min}$ .  
(Note that  $(A \cap R^S) \setminus R_{\min}$  could be empty.)

For  $i = k$  to  $1$ , if  $-\epsilon\Delta \leq d(A \setminus \{\rho_i, \rho_{i+1}, \dots, \rho_k\}) \leq \epsilon\Delta$ , then let  $A = A \setminus \{\rho_i, \rho_{i+1}, \dots, \rho_k\}$ .
    - (c) If  $c(A) < c(M_{\text{best}})$ , then let  $M_{\text{best}} = M_A$ .
  5. If  $M_{\text{best}} \neq \text{NULL}$ , then output  $M_{\text{best}}$ , otherwise output “None” and halt.
-



**THEOREM 4.3.** *There is a  $(2 - (\epsilon - \delta)/(2 + 3\epsilon))$ -approximation algorithm for MinESE whose running time is  $O(n^{3+2(\frac{1+\epsilon}{\delta})})$  for an arbitrary  $\delta$  such that  $0 < \delta < \epsilon$ .*

**PROOF. Correctness Proof.** Clearly, if there is no stable matching  $M$  such that  $-\epsilon\Delta \leq d(M) \leq \epsilon\Delta$ , then the algorithm answers “None.” Otherwise, suppose that there is a feasible solution, and let  $M_{opt}$  be an optimal solution. We first show that Algorithm 2 finds a feasible solution. Let  $OPT$  be the rotation set corresponding to  $M_{opt}$ , and define  $OPT^L = OPT \cap R^L$ . Then,  $d(OPT^L) \leq d(OPT) = d(M_{opt}) \leq \epsilon\Delta$ . Because  $w_d(\rho) > \delta\Delta$  for any rotation  $\rho \in OPT^L$ ,  $|OPT^L| < \frac{d(OPT^L) - d(M_0)}{\delta\Delta} \leq \frac{|d(M_0)| + \epsilon\Delta}{\delta\Delta} = \frac{1+\epsilon}{\delta}$ . This means Algorithm 2 selects  $OPT^L$  at Step 4 as  $R$ , and we consider this particular execution of Step 4. We show that in this execution, Algorithm 2 finds a feasible solution. Let  $A_{opt} = \text{minEgalitarian}(OPT^L, R^L \setminus OPT^L)$ . There are three cases:

- (i)  $-\epsilon\Delta \leq d(A_{opt}) \leq \epsilon\Delta$ .  $M_{A_{opt}}$  is clearly a feasible solution.
- (ii)  $d(A_{opt}) < -\epsilon\Delta$ . Note that  $A_{opt} \cup \{\rho_1, \rho_2, \dots, \rho_k\}$  is the maximal closed subset that contains exactly the same elements from  $R^L$  as  $OPT$ . Hence  $OPT \subseteq A_{opt} \cup \{\rho_1, \rho_2, \dots, \rho_k\}$ , and so,  $d(A_{opt} \cup \{\rho_1, \rho_2, \dots, \rho_k\}) \geq d(M_{opt}) \geq -\epsilon\Delta$ . Note also that any rotation  $\rho_i$  ( $1 \leq i \leq k$ ) satisfies  $w_d(\rho_i) \leq \delta\Delta$ . Hence there must be  $j$  ( $1 \leq j \leq k$ ) such that  $-\epsilon\Delta \leq d(A_{opt} \cup \{\rho_1, \rho_2, \dots, \rho_j\}) \leq -(\epsilon - \delta)\Delta$ . (It can be easily seen that  $A_{opt} \cup \{\rho_1, \rho_2, \dots, \rho_j\}$  is closed.)
- (iii)  $d(A_{opt}) > \epsilon\Delta$ . Note that  $A_{opt} \setminus \{\rho_1, \rho_2, \dots, \rho_k\}$  is the minimal closed subset that contains exactly the same elements from  $R^L$  as  $OPT$ . Hence  $A_{opt} \setminus \{\rho_1, \rho_2, \dots, \rho_k\} \subseteq OPT$ , and so,  $d(A_{opt} \setminus \{\rho_1, \rho_2, \dots, \rho_k\}) \leq d(M_{opt}) \leq \epsilon\Delta$ . Note also that any rotation  $\rho_i$  ( $1 \leq i \leq k$ ) satisfies  $w_d(\rho_i) \leq \delta\Delta$ . Hence there must be  $j$  ( $1 \leq j \leq k$ ) such that  $(\epsilon - \delta)\Delta \leq d(A_{opt} \setminus \{\rho_j, \rho_{j+1}, \dots, \rho_k\}) \leq \epsilon\Delta$ . (It can be easily seen that  $A_{opt} \setminus \{\rho_j, \rho_{j+1}, \dots, \rho_k\}$  is closed.)

Next, we analyze the approximation ratio. Let  $M^*$  be the matching found in this particular execution of Step 4. We show that  $c(M^*) \leq (2 - (\epsilon - \delta)/(2 + 3\epsilon))c(M_{opt})$ , which gives us a proof for the approximation ratio. We first prove the following two lemmas:

**LEMMA 4.4.** *For any stable matching  $M$ ,  $|d(M)| < c(M)$ .*

**PROOF.** If  $d(M) \geq 0$ , then  $c(M) - |d(M)| = 2 \sum_{(m,w) \in M} p_w(m) > 0$ . Otherwise,  $c(M) - |d(M)| = 2 \sum_{(m,w) \in M} p_m(w) > 0$ .  $\square$

**LEMMA 4.5.** *Let  $R = \{\rho_1, \dots, \rho_{r-1}\}$  be a set of rotations and let  $M_1, \dots, M_r$  be stable matchings such that  $M_{i+1} = M_i / \rho_i$  for  $1 \leq i < r$ . Then,  $|c(M_r) - c(M_1)| < d(M_r) - d(M_1)$ .*

**PROOF.** Suppose that for a pair  $(m, w) \in M_i$ ,  $m$  and  $w$  are included in a rotation  $\rho_i$ . Let  $m' = M_{i+1}(w)$  and  $w' = M_{i+1}(m)$ . By the properties of the rotation [Gusfield and Irving 1989],  $m$  prefers  $w$  to  $w'$  and  $w$  prefers  $m'$  to  $m$ . Let  $d(m) = p_m(w') - p_m(w)$  and  $d(w) = p_w(m) - p_w(m')$ . Then  $d(m) > 0$  and  $d(w) > 0$ , and it follows that

$$|c(M_{i+1}) - c(M_i)| = \left| \sum_m d(m) - \sum_w d(w) \right| < \left| \sum_m d(m) + \sum_w d(w) \right| = d(M_{i+1}) - d(M_i),$$

where the summations of  $m$  ( $w$ ) range over the set of men (women) that have different partners in  $M_i$  and  $M_{i+1}$ . By summing up the above inequality for all  $i$ , we have

$$|c(M_r) - c(M_1)| \leq \sum_{i=1}^{r-1} |c(M_{i+1}) - c(M_i)| < \sum_{i=1}^{r-1} (d(M_{i+1}) - d(M_i)) = d(M_r) - d(M_1).$$

□

Note that  $A_{opt} = \min\text{Egalitarian}(OPT^L, R^L \setminus OPT^L)$ . Therefore,  $c(A_{opt}) \leq c(M_{opt})$  since  $OPT$ , the rotation set corresponding to  $M_{opt}$ , is one of the candidates for  $A_{opt}$ . We will consider the following four cases (note that  $d(A_{opt}) \geq -\Delta$  for any stable matching  $M$ ):

**Case (i):  $-\epsilon\Delta \leq d(A_{opt}) \leq \epsilon\Delta$ .** In this case,  $M^* = M_{A_{opt}}$ , which is an optimal solution since  $c(A_{opt}) = c(M_{opt})$ .

**Case (ii):  $-\Delta \leq d(A_{opt}) < -\epsilon\Delta$ .** In this case, the “If” part of Step 4(b) of Algorithm 2 is executed. The analysis can be done in a similar way as we have seen at the beginning of this section (Fig. 3). By Lemma 4.5,  $|c(M^*) - c(A_{opt})| < d(M^*) - d(A_{opt})$ . If  $c(M^*) - c(A_{opt}) \leq 0$ , then  $c(M^*) \leq c(A_{opt}) \leq c(M_{opt})$ , which implies that  $M^*$  is optimal. Therefore, in the following, we consider the case  $c(M^*) - c(A_{opt}) > 0$ . By Lemma 4.4,  $|d(A_{opt})| < c(A_{opt})$ . Note also that  $c(A_{opt}) \leq c(M_{opt})$  and  $d(A_{opt}) < 0$ . So,  $-d(A_{opt}) < c(A_{opt}) \leq c(M_{opt})$ . Furthermore,  $d(M^*) < -(\epsilon - \delta)\Delta \leq (\epsilon - \delta)d(A_{opt})$ . Putting these facts together, we have  $c(M^*) < (2 - (\epsilon - \delta))c(M_{opt}) < (2 - (\epsilon - \delta)/(2 + 3\epsilon))c(M_{opt})$ .

**Case (iii):  $\epsilon\Delta < d(A_{opt}) \leq (2 + 3\epsilon)\Delta$ .** In this case, the “Else if” part of Step 4(b) of Algorithm 2 is executed. We have  $|c(A_{opt}) - c(M^*)| < d(A_{opt}) - d(M^*)$  by Lemma 4.5. Since  $d(M^*) \geq (\epsilon - \delta)\Delta$  and (iii) hold,  $|c(A_{opt}) - c(M^*)| < (1 - (\epsilon - \delta)/(2 + 3\epsilon))d(A_{opt})$ . If  $c(M^*) - c(A_{opt}) \leq 0$ , then  $M^*$  is optimal as discussed in Case (ii). Therefore, suppose that  $c(M^*) - c(A_{opt}) > 0$ . Since  $|d(A_{opt})| < c(A_{opt})$  by Lemma 4.4 and  $c(A_{opt}) \leq c(M_{opt})$ ,  $c(M^*) < (2 - (\epsilon - \delta)/(2 + 3\epsilon))c(M_{opt})$ .

**Case (iv):  $(2 + 3\epsilon)\Delta < d(A_{opt})$ .** Since both  $M_{opt}$  and  $M^*$  can be obtained by repeatedly eliminating rotations from  $M_0$ ,  $|c(M_{opt}) - c(M_0)| < d(M_{opt}) - d(M_0)$  and  $|c(M^*) - c(M_0)| < d(M^*) - d(M_0)$  by Lemma 4.5 (See Fig. 4). Since both  $d(M_{opt})$  and  $d(M^*)$  are at most  $\epsilon\Delta$ ,  $c(M^*) - c(M_{opt}) \leq 2(1 + \epsilon)\Delta$  (note that  $|d(M_0)| = \Delta$ ). It follows that  $c(M^*) - c(M_{opt}) \leq 2(1 + \epsilon)d(A_{opt})/(2 + 3\epsilon) = (1 - \epsilon/(2 + 3\epsilon))d(A_{opt})$ . Since we have  $|d(A_{opt})| < c(A_{opt})$  by Lemma 4.4 and  $c(A_{opt}) \leq c(M_{opt})$ ,  $c(M^*) < (2 - \epsilon/(2 + 3\epsilon))c(M_{opt})$ .

**Time Complexity.** Steps 1, 2, 3, and 5 can be executed in  $O(n^2)$  time. Step 4(a) is performed in time  $O(n^3)$  by Proposition 4.2. We can see that Steps 4(b) and 4(c) can be performed in time  $O(n^2)$  by a similar analysis of Algorithm 1. The number of repetitions of Step 4 can be analyzed in the same way as for the proof of Theorem 3.1, which is  $O(n^{2(\frac{1+\epsilon}{\delta})})$ . Hence the time complexity of Algorithm 2 is  $O(n^{3+2(\frac{1+\epsilon}{\delta})})$ .

## 5. CONCLUDING REMARKS

In this paper, we gave a polynomial time algorithm for finding near optimal sex-equal stable matching. Furthermore, we proved strongly NP-hardness and devel-

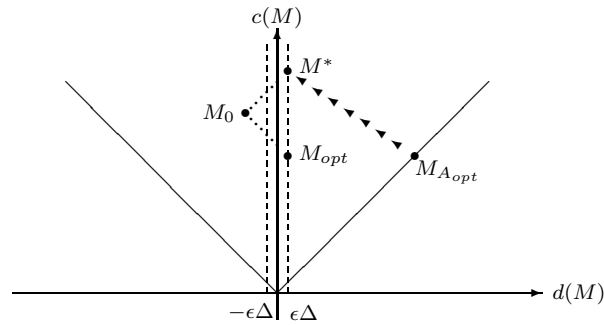


Fig. 4. Case (iv)

oped a polynomial time approximation algorithm whose approximation ratio is less than 2 for MinESE. Our future work is to improve the approximation ratio of MinESE.

#### REFERENCES

- BONSMAN, P. S. 2007. Most balanced minimum cuts and partially ordered knapsack. In *Proc. of the 6th Cologne-Twente Workshop on Graphs and Combinatorial Optimization*. 17–21.
- CHARIKAR, M. AND WIRTH, A. 2004. Maximizing quadratic programs: extending Grothendieck's inequality. In *Proc. of the 45th Annual Symposium on Foundations of Computer Science*. 54–60.
- FEDER, T. 1992. A new fixed point approach for stable networks and stable marriages. *Journal of Computer and System Sciences* 45, 233–284.
- FEDER, T. 1994. Network flow and 2-satisfiability. *Algorithmica* 11, 291–319.
- GALE, D. AND SHAPLEY, L. S. 1962. College admissions and the stability of marriage. *Amer. Math. Monthly* 69, 9–15.
- GUSFIELD, D. 1987. Three fast algorithms for four problems in stable marriage. *SIAM J. Comput.* 16, 1, 111–128.
- GUSFIELD, D. AND IRVING, R. W. 1989. *The Stable Marriage Problem: Structure and Algorithms*. MIT Press, Boston, MA.
- GUSFIELD, D., IRVING, R. W., LEATHER, P., AND SAKS, M. 1987. Every finite distributive lattice is a set of stable matchings for a small instance. *J. Combin. Theory A* 44, 304–309.
- HALLDÓRSSON, M. M., IRVING, R. W., IWAMA, K., MANLOVE, D. F., MIYAZAKI, S., MORITA, Y., AND SCOTT, S. 2003. Approximability results for stable marriage problems with ties. *Theoretical Computer Science* 306, 431–447.
- IRVING, R. W. AND LEATHER, P. 1986. The complexity of counting stable marriages. *SIAM J. Comput.* 15, 3, 655–667.
- IRVING, R. W., LEATHER, P., AND GUSFIELD, D. 1987. An efficient algorithm for the “optimal” stable marriage. *J. ACM* 34, 3, 532–543.
- JOHNSON, D. S. AND NIEMI, K. A. 1983. On knapsacks, partitions, and a new dynamic programming technique for trees. *Mathematics of Operations Research* 8, 1, 1–14.
- KATO, A. 1993. Complexity of the sex-equal stable marriage problem. *Japan Journal of Industrial and Applied Mathematics (JJIAM)* 10, 1–19.
- KOLLIPOPOULOS, S. G. AND STEINER, G. 2007. Partially ordered knapsack and applications to scheduling. *Discrete Applied Mathematics* 155, 8, 889–897.
- NESTEROV, Y. 1998. Semidefinite relaxation and nonconvex quadratic optimization. *Optimization Methods and Software* 9, 141–160.
- ROMERO-MEDINA, A. 2001. ‘Sex-equal’ stable matchings. *Theory and Decision* 50, 197–212.

RECEIVED ...; REVISED ...; ACCEPTED ...